1 The Mathematical Framework

It is the goal of quantum theory – just as of every other physical theory – to predict the results of experiments and to justify these predictions. To this end, it is necessary to describe the state of the physical system at the beginning of an experiment. One must also be able to formulate the evolution of the system under external influences and to predict the effect of its interaction with the measurement apparatus. The mathematical framework which has proven most expedient for the formulation of quantum mechanics is the theory of the Hilbert space and probability theory. The fundamental connection between mathematical quantities and physical reality is established by the following associations:

Quantum system	$\leftrightarrow$	Hilbert space.
Quantum state	$\leftrightarrow$	vector in (or, more generally: density operator on) the Hilbert space.
Evolution of the quantum state	$\leftrightarrow$	linear operators, which act on the vectors, or linear operators, which act on the operator space (Liouville space).
Predictions	$\leftrightarrow$	probabilistic statements.

We will describe this basic scheme of the quantum theory in detail. In this chapter, we first collect the required mathematical definitions and theorems. We shall not prove all of the mathematical theorems; in particular, we assume that the reader has already had some contact with quantum theory, so that our treatment here can be brief.

Since we will be concerned exclusively with d-level quantum systems ($d = 2, 3, \ldots$), we make use of a restriction which will greatly simplify our treatment:

> ***General mathematical assumption*:** *We consider only quantum systems which can be described with the aid of a* finite-dimensional *Hilbert space $\mathcal{H}_d$ of dimension $d = 2, 3, \ldots$.*

This restriction is justified, since the essential conceptual problems and the new ideas and central methods can all be introduced by referring to a finite-dimensional Hilbert space. We wish to avoid the addition of mathematical subtleties to the conceptual physical problems. For the majority of physically-relevant cases which require a description in infinite-dimensional Hilbert space, the results for finite-dimensional spaces can be directly applied.

As is usual in theoretical physics, we will make use of the *Dirac notation*. In this framework, it is expedient to place the dyadic decomposition of operators at the centre of our treatment. This is important for practical applications, since it permits a simple, direct reading-off of the properties and effects of the operators.

Entangled Systems: New Directions in Quantum Physics. Jürgen Audretsch

ISBN: 978-3-527-40684-5

1.1 Hilbert Vector Space

1.1.1 The Scalar Product and the Dirac Notation

A d-dimensional Hilbert space $\mathcal{H}_d$ is a linear complex vector space in which a scalar product is defined. The vectors are denoted by $|\varphi\rangle$, $|\psi\rangle$, $|u\rangle$, $|\Phi\rangle$, etc.; $|\text{null}\rangle$ is the null vector.

Addition and multiplication with a complex number, linear independence, the basis and dimensionality of the Hilbert space $\mathcal{H}_d$ are defined in an analogous way to the corresponding concepts in real vector spaces.

A complex number is associated to a pair of vectors $|\varphi\rangle$ and $|\psi\rangle$ as their *scalar product* or *inner product*, which we write in the form $\langle\varphi|\psi\rangle$. As the basis of this *Dirac notation*[1], we have introduced a *ket space* with the *ket vectors* $|\varphi\rangle, |\psi\rangle, \ldots$ and its dual vector space of the *bra vectors* $\langle\chi|, \langle\theta|, \ldots$ (space of linear functionals). There is a declared correspondence between the vectors of the ket space and of the bra space,

$$|\varphi\rangle \overset{d.c.}{\leftrightarrow} \langle\varphi| \,, \tag{1.1}$$

which is called the *dual correspondence* for vectors. We use the same central symbol as an expression of the dual correspondence. Here, a ket vector $|\varphi\rangle = c_1|\varphi_1\rangle + c_2|\varphi_2\rangle$ is associated via a one-to-one correspondence with a bra vector $\langle\varphi| = c_1^*\langle\varphi_1| + c_2^*\langle\varphi_2|$ ($*$ signifies the complex conjugate). The ordering within the product $\langle\varphi|\psi\rangle$ is thus important. We have:

$$\begin{aligned}
&\langle\varphi|\psi\rangle = \langle\psi|\varphi\rangle^* \\
&\langle\varphi|c_1\psi_1 + c_2\psi_2\rangle = c_1\langle\varphi|\psi_1\rangle + c_2\langle\varphi|\psi_2\rangle \,, \quad c_1, c_2 \in \mathbb{C} \\
&\langle\varphi|\varphi\rangle \geq 0 \,\forall\, |\varphi\rangle \in \mathcal{H}_n \,, (\langle\varphi|\varphi\rangle = 0 \Leftrightarrow |\varphi\rangle = |\text{null}\rangle) \,.
\end{aligned} \tag{1.2}$$

From this, it follows that

$$\langle c_1\varphi_1 + c_2\varphi_2|\psi\rangle = c_1^*\langle\varphi_1|\psi\rangle + c_2^*\langle\varphi_2|\psi\rangle \,. \tag{1.3}$$

The scalar product is linear in its second argument and *antilinear* in its first argument. When $\langle\varphi|\psi\rangle = 0$ holds, the vectors are termed *orthogonal* to each other.

The product induces a *norm* on the Hilbert space according to

$$\|\varphi\| =: \||\varphi\rangle\| := \sqrt{\langle\varphi|\varphi\rangle} \,. \tag{1.4}$$

It vanishes if and only if $|\varphi\rangle$ is the zero vector. We mention without proof *Schwarz's inequality*

$$|\langle\varphi|\psi\rangle| \;\leq\; \|\varphi\|\,\|\psi\| \tag{1.5}$$

and the *triangle relations*

$$\|\varphi\| - \|\psi\| \;\leq\; \|\psi - \varphi\| \;, \qquad \|\varphi + \psi\| \;\leq\; \|\varphi\| + \|\psi\| \;. \tag{1.6}$$

[1] Following Dirac, the scalar product is written as $\langle\varphi|\psi\rangle$ and called a "bracket". Its components "bra" $\langle\varphi|$ and "ket" $|\psi\rangle$ denote independent vectors

One can show by substitution that

$$\langle\varphi|\psi\rangle = \frac{1}{4}\left(\|\varphi+\psi\|^2 - \|\varphi-\psi\|^2 + i\,\|\varphi-i\psi\|^2 - i\,\|\varphi+i\psi\|^2\right) \tag{1.7}$$

holds, as well as the *parallelogram equation*

$$\|\varphi+\psi\|^2 + \|\varphi-\psi\|^2 = 2\,\|\varphi\|^2 + 2\,\|\psi\|^2 \ . \tag{1.8}$$

For a set of vectors $\{|\varphi_1\rangle, |\varphi_2\rangle, \ldots, |\varphi_l\rangle\}$ in $\mathcal{H}_d$, span$(|\varphi_1\rangle, \ldots, |\varphi_l\rangle)$ denotes the set of all possible linear combinations of these vectors. This set forms a subspace of $\mathcal{H}_d$ which is itself a Hilbert space. We denote an *orthonormal basis* by *ONB*. For an ONB $\{|i\rangle, i = 1, \ldots, d\}$, the decomposition

$$|\varphi\rangle = \sum_{i=1}^{d} |i\rangle\langle i|\varphi\rangle \tag{1.9}$$

applies, defining the *components* $\langle i|\varphi\rangle$ of the vector $|\varphi\rangle$ with respect to the ONB. The set of all vectors $|\psi\rangle$ which are orthogonal to all the vectors in a subspace $\hat{\mathcal{H}}$ of $\mathcal{H}$ forms an additional subspace of $\mathcal{H}$, which is called the *orthogonal complement* $\hat{\mathcal{H}}^\perp$. The direct sum of the two subspaces is again the Hilbert space, $\mathcal{H} = \hat{\mathcal{H}} \oplus \hat{\mathcal{H}}^\perp := \{\alpha|\chi\rangle + \beta|\psi\rangle$ with $|\chi\rangle \in \hat{\mathcal{H}},\ |\psi\rangle \in \hat{\mathcal{H}}^\perp$ and $\alpha, \beta \in \mathbb{C}\}$.

1.1.2 Linear Operators on the Hilbert Space

Linear operators $A, B, \ldots$ map ket vectors in a linear way onto one another

$$\begin{aligned}
A(\alpha|\psi\rangle + \beta|\phi\rangle) &= \alpha A|\psi\rangle + \beta A|\phi\rangle && \text{linearity}(\alpha, \beta \in \mathbb{C}) \\
(A+B)|\psi\rangle &= A|\psi\rangle + B|\psi\rangle && \text{sum} \\
(AB)|\psi\rangle &= A(B|\psi\rangle) && \text{product} \\
A|\psi_a\rangle &= a|\psi_a\rangle && \textit{eigenvector } |\psi_a\rangle \text{ of } A \\
& && \textit{eigenvalue } a \text{ of } A \\
\mathbb{1}|\psi\rangle &= |\psi\rangle && \textit{identity operator, unit operator}.
\end{aligned} \tag{1.10}$$

The domain of definition of A need not be the entire Hilbert space, and its co-domain need not be identical with its definition range. When necessary, we will make a remark on this point. For the *inverse operator* A^{-1}, we have $AA^{-1} = A^{-1}A = \mathbb{1}$. We wish to extend the Dirac notation further, and therefore adopt the convention that operators on the bra space (arrow to the left) act from the right on bra vectors:

$$\langle\varphi'| = \langle\varphi|\overleftarrow{B}\ . \tag{1.11}$$

The operators on the ket space (arrow to the right) act correspondingly from the left. For the resulting vector, we write

$$|\psi'\rangle = \overrightarrow{A}|\psi\rangle =: |A\psi\rangle\,. \tag{1.12}$$

Via the dual correspondence (1.1), a ket vector $|A\psi\rangle$ corresponds to a bra vector $\langle A\psi|$:

$$|A\psi\rangle \overset{d.c.}{\leftrightarrow} \langle A\psi| \; . \tag{1.13}$$

We in addition introduce a dual correspondence for the operators. Referring to the Dirac notation, the bra operator which corresponds to a ket operator $\overrightarrow{A}$ is likewise denoted by the same central symbol A:

$$\overrightarrow{A} \overset{d.c.}{\leftrightarrow} \overleftarrow{A} \tag{1.14}$$

The correspondence is determined by the following condition on the scalar products (first equation):

$$(\langle\varphi|\overleftarrow{A})|\psi\rangle = \langle\varphi|(\overrightarrow{A}|\psi\rangle) =: \langle\varphi|A|\psi\rangle \; . \tag{1.15}$$

The second equation is an abbreviation, with the compactness which is characteristic of the Dirac notation. Furthermore, we write $A|\psi\rangle$ for $\overrightarrow{A}|\psi\rangle$ and $\langle\psi|A$ for $\langle\psi|\overleftarrow{A}$.

Adjoint operators The dual correspondence leads from $|\psi\rangle$ to $\langle\psi|$. By application of $\overrightarrow{A}$ to $|\psi\rangle$, we obtain $|A\psi\rangle$, and the dual correspondence (1.13) defines $\langle A\psi|$. One can however also obtain $\langle A\psi|$ directly in bra space by applying an operator $\overleftarrow{A}^\dagger$ to $\langle\psi|$:

$$\langle A\psi| =: \langle\psi|\overleftarrow{A}^\dagger \; . \tag{1.16}$$

The operator $\overleftarrow{A}^\dagger$ thus defined is called the *adjoint operator* to $\overleftarrow{A}$. By $\overrightarrow{A}$, both $\overleftarrow{A}$ as well as $\overleftarrow{A}^\dagger$ are defined in the bra space. Finally, via the dual correspondence, the adjoint operator $\overrightarrow{A}^\dagger$ in the ket space is:

$$\overrightarrow{A}^\dagger \overset{d.c.}{\leftrightarrow} \overleftarrow{A}^\dagger \; . \tag{1.17}$$

In the Dirac notation, we can leave off the arrows as in Eq. (1.15) and thereby omit the explicit reference to the two spaces. We evaluate Eq. (1.16) using Eq. (1.15):

$$\langle A\psi|\varphi\rangle = (\langle\psi|\overleftarrow{A}^\dagger)|\varphi\rangle = \langle\psi|(\overrightarrow{A}^\dagger|\varphi\rangle) = \langle\psi|A^\dagger\varphi\rangle = \langle\psi|A^\dagger|\varphi\rangle \tag{1.18}$$

and summarise the result:

$$\langle\overleftarrow{A}\psi|\varphi\rangle = \langle\psi|\overrightarrow{A}^\dagger\varphi\rangle = \langle\psi|A^\dagger|\varphi\rangle \; . \tag{1.19}$$

With $\langle\varphi|\psi\rangle = \langle\psi|\varphi\rangle^*$, it follows from Eq. (1.19) that

$$\langle\psi|A^\dagger|\varphi\rangle = \langle\varphi|A\psi\rangle^* = \langle\varphi|A|\psi\rangle^* \; . \tag{1.20}$$

Repeated application of Eq. (1.20) yields

$$\langle\varphi|A|\psi\rangle = (\langle\varphi|A|\psi\rangle^*)^* = \langle\psi|A^\dagger|\varphi\rangle^* = \langle\varphi|(A^\dagger)^\dagger|\psi\rangle \tag{1.21}$$

for arbitrary vectors $\langle\varphi|$ and $\langle\psi|$. Thus, we find

$$(A^\dagger)^\dagger = A \tag{1.22}$$

and we obtain the corresponding relation to Eq. (1.19)

$$\langle A^\dagger\psi|\varphi\rangle = \langle\psi|A\varphi\rangle = \langle\psi|A|\varphi\rangle \; . \tag{1.23}$$

In a similar manner, one can readily convince oneself of the validity of the following operator relations:

$$\left(A^{-1}\right)^\dagger = \left(A^\dagger\right)^{-1} \; , \quad (cA)^\dagger = c^* A^\dagger \tag{1.24}$$

$$(A+B)^\dagger = A^\dagger + B^\dagger \, , \quad (AB)^\dagger = B^\dagger A^\dagger \, . \tag{1.25}$$

In addition to the definition (1.16), equations (1.22) and (1.23) are frequently used.

Dyadic decomposition We introduce the *dyadic product* (outer product) or the *dyad* $|u\rangle\langle v|$ of two vectors $|u\rangle$ and $|v\rangle$. It is a linear operator

$$|\varphi\rangle \rightarrow |\psi\rangle = (|u\rangle\langle v|)|\varphi\rangle = |u\rangle\langle v|\varphi\rangle \, , \tag{1.26}$$

which produces a vector parallel to $|u\rangle$. The scalar product $\langle v|\varphi\rangle$ appears as a complex factor. $|u\rangle\langle v|$ is in the first instance to be considered as an overall symbol which cannot be decomposed, denoting a linear operator with certain properties. These include:

$$(\alpha|u\rangle\langle v|)^\dagger = \alpha^*|v\rangle\langle u| \, . \tag{1.27}$$

For operator products, we find

$$A|u\rangle\langle v| = |Au\rangle\langle v| \, , \qquad\qquad |u\rangle\langle v|A = |u\rangle\langle A^\dagger v| \; . \tag{1.28}$$

As we have seen in Eq. (1.9), the identity operator can be represented in terms of a dyad with the aid of an ONB $\{|i\rangle, i = 1, \ldots, d\}$ of the Hilbert space:

$$\mathbb{1} = \sum_i |i\rangle\langle i| \; . \tag{1.29}$$

This is also referred to as the *completeness relation* or the *dyadic decomposition of the identity operator*. *From Eq. (1.29), it follows using Eq. (1.26) that every linear operator has a* dyadic decomposition *(outer product representation)*

$$A = \sum_{i,j} |i\rangle\langle i|A|j\rangle\langle j| = \sum_{i,j} \langle i|A|j\rangle|i\rangle\langle j| = \sum_{i,j} A_{ij}|i\rangle\langle j| \tag{1.30}$$

with the matrix elements $A_{ij} := \langle i|A|j\rangle$. We can read off the equations (1.26) through (1.30) as a suggestive mnemonic rule, *that* $|u\rangle\langle v|$ *and the dyadic decomposition of* A *can be taken formally to act in such a way, as if* $|u\rangle$ *and* $\langle v|$ *were independent vectors and not parts of*

an overall symbol $|u\rangle\langle v|$. This is one of the great advantages of the Dirac notation. For the adjoint operator, we obtain

$$A^\dagger = \sum_{i,j} A^*_{ij} |j\rangle\langle i| \; . \tag{1.31}$$

Via the *supremum norm* $\|A\|$, one can associate to a linear operator A a positive number:

$$\|A\| := \sup_{\langle\varphi|\varphi\rangle=1} |\langle\varphi|A|\varphi\rangle| \, . \tag{1.32}$$

Trace The *trace* is a frequently-used complex-valued function of a linear operator:

$$\mathrm{tr}[A] := \sum_i \langle i|A|i\rangle = \sum_i A_{ii} \, , \quad \{|i\rangle\} \text{ ONB} \; . \tag{1.33}$$

The trace of an operator is independent of the choice of the basis. The proof of this statement demonstrates the usefulness of the dyadic decomposition (1.29) of the identity operator. Let $\{|l_i\rangle\}$ and $\{|m_j\rangle\}$ be an arbitrary ONB; then using the mnemonic rule above, we find:

$$\begin{aligned} \mathrm{tr}[A] &= \sum_i \langle l_i|A|l_i\rangle = \sum_{i,j,k} \langle l_i|m_j\rangle\langle m_j|A|m_k\rangle\langle m_k|l_i\rangle \\ &= \sum_{i,j,k} \langle m_k|l_i\rangle\langle l_i|m_j\rangle\langle m_j|A|m_k\rangle = \sum_{j,k} \langle m_k|m_j\rangle\langle m_j|A|m_k\rangle \\ &= \sum_j \langle m_j|A|m_j\rangle \; . \end{aligned} \tag{1.34}$$

In a similar manner, using Eq. (1.29) one can prove the following properties of the trace:

$$\begin{aligned} \mathrm{tr}[AB] &= \mathrm{tr}[BA] && \text{cyclic permutations} \\ \mathrm{tr}[A+B] &= \mathrm{tr}[A] + \mathrm{tr}[B] && \text{linearity} \\ \mathrm{tr}[\alpha A] &= \alpha \, \mathrm{tr}[A] && \text{linearity} \\ \mathrm{tr}[A|\psi\rangle\langle\psi|] &= \langle\psi|A|\psi\rangle && \text{expectation value of } A \\ \mathrm{tr}[|\varphi\rangle\langle\psi|] &= \langle\psi|\varphi\rangle && \text{trace of a dyad} \\ \mathrm{tr}[A^\dagger] &= (\mathrm{tr}[A])^* && \text{adjoint operator} \end{aligned} \tag{1.35}$$

The terms *expectation value* or *mean value* of A used in quantum physics will later be justified on the basis of physical arguments.

1.1.3 Normal Operators and Spectral Decompositions

Among the linear operators on $\mathcal{H}_d$, those which are diagonalisable, also called the *normal operators*, play a particularly important role in mathematics and physics. An operator N is termed *diagonalisable* if there exists an ONB $\{|i\rangle\}$ of $\mathcal{H}_d$ and a set of complex numbers $\lambda_i \in \mathbb{C}$ such that

$$N|i\rangle = \lambda_i |i\rangle \tag{1.36}$$

holds. Here, $\lambda_i = 0$ is not excluded. An immediate result is that the matrix of N in the ONB of the eigenvectors is diagonal

$$N_{ij} = \langle i|N|j\rangle = \lambda_i \delta_{ij} \tag{1.37}$$

and therefore the operator N can be written in the form of a *spectral decomposition*

$$N = \sum_i \lambda_i |i\rangle\langle i| \,, \quad \lambda_i \in \mathbb{C} \,. \tag{1.38}$$

This is also called the *orthogonal decomposition*. The ONB $\{|i\rangle\}$ of Eq. (1.36) is also referred to as the *eigenbasis* of N. Conversely, the diagonalisability condition (1.36) follows directly from each of these relations. If there are $g \geq 2$ linearly-independent eigenvectors $|j_l\rangle$ belonging to an eigenvalue λ_j of the eigenvalue problem (1.36), where l = 1... g, then λ_j is said to be *g-fold degenerate*. Every linear combination of these eigenvectors

$$|\psi\rangle = \sum_{l=1}^{g} c_l |j_l\rangle \tag{1.39}$$

is then likewise an eigenvector belonging to the eigenvalue λ_j. The eigenvectors span a g-dimensional subspace $\mathcal{H}_{(j)}$ of $\mathcal{H}$. The *projector*

$$P = \sum_{l=1}^{g} |j_l\rangle\langle j_l| \,, \quad P^\dagger = P \,; \quad P^2 = P \,, \tag{1.40}$$

projects into the subspace $\mathcal{H}_{(j)}$. The projector $Q = 1 - P$ projects into the orthogonal complement of $\mathcal{H}_{(j)}$, i.e. $\mathcal{H}_{(j)}^{\perp}$. The subspaces belonging to different eigenvalues are orthogonal to one another.

Diagonalisability is by no means a trivially-occurring property. Even in the two-dimensional Hilbert space $\mathcal{H}_2$, there are frequently-used operators which are not diagonalisable. An example is

$$A = |0\rangle\langle 1| \quad \text{with} \quad \langle 0|1\rangle = 0 \quad \text{and} \quad \langle 0|0\rangle = \langle 1|1\rangle = 1 \tag{1.41}$$

as can be shown with the help of the following theorem.

In order to recognise whether a given operator is a normal operator, the following central theorem is very useful: *A necessary and sufficient condition that an operator N can be spectrally decomposed – that is, it is diagonalisable – is the vanishing of the commutator $([A, B]_- := AB - BA)$ of N and $N^\dagger$:*

$$[N, N^\dagger]_- = 0 \,. \tag{1.42}$$

The proof of this theorem can serve as an example of the application of the formalism which we have thus far constructed. The fact that diagonalisability follows from Eq. (1.42) is clear. The converse direction of the proof can be divided into two steps:

1. Step: Each operator in $\mathcal{H}_d$ has at least one eigenvalue λ and one eigenvector $|1\rangle$, which can be found with the aid of the secular equation:

$$N|1\rangle = \lambda|1\rangle\,, \quad \langle 1|N^\dagger = \lambda^*\langle 1|\,. \tag{1.43}$$

It then follows that

$$\langle 1|N|1\rangle = \lambda\,, \quad \langle 1|N^\dagger|1\rangle = \lambda^* \tag{1.44}$$

and thus

$$N^\dagger|1\rangle = \lambda^*|1\rangle + |a\rangle\,, \quad \langle 1|N = \lambda\langle 1| + \langle a| \tag{1.45}$$

with $\langle 1|a\rangle = 0$. Using the normality condition $[N, N^\dagger]_- = 0$, we find after evaluation with Eqs. (1.43) and (1.45)

$$0 = \langle 1|[N, N^\dagger]_-|1\rangle = \langle a|a\rangle\,. \tag{1.46}$$

$|a\rangle$ is thus the null vector $|\text{null}\rangle$ and (1.45) can be written as follows:

$$N^\dagger|1\rangle = \lambda^*|1\rangle\,, \quad \langle 1|N = \lambda\langle 1|\,. \tag{1.47}$$

We have thus determined the action of N and $N^\dagger$ on $|1\rangle$.

2. Step: We complete $|1\rangle$ to obtain an ONB $\{|i\rangle\}$ and introduce with the aid of the dual notation for N:

$$N = \sum_{ij} n_{ij}|i\rangle\langle j|, \quad n_{ij} := \langle i|N|j\rangle, \quad n_{1i} = n_{i1} = \lambda\delta_{i1} \tag{1.48}$$

the operator M

$$M := N - \lambda|1\rangle\langle 1|\,, \quad M = \sum_{i,j\neq 1} n_{ij}|i\rangle\langle j|\,. \tag{1.49}$$

M is the restriction of N to the orthogonal complement of $|1\rangle$.

Making use of Eqs. (1.43) and (1.47), we can show that M is also a normal operator, $([M, M^\dagger]_- = 0)$. The same procedure can be applied to it in the subspace which is orthogonal to $|1\rangle$. M also has an eigenvector, which we denote as $|2\rangle$. We now complete $|1\rangle$ and $|2\rangle$ to an ONB and repeat the procedure. We continue in the same manner until the entire Hilbert space is used up and $|1\rangle$ has been completed to a well-defined ONB. At the same time, N has been spectrally decomposed with respect to this basis. This concludes the proof. The fact that the operator A of Eq. (1.41) does *not* fulfill the condition (1.42) can be readily verified.

The diagram in Fig. 1.1 demonstrates how the various properties of the operators in Hilbert space correspond in an intuitively clear way to an increasing specialisation in the dyadic decomposition. In the following section, we will go through this diagram step by step from above to below.

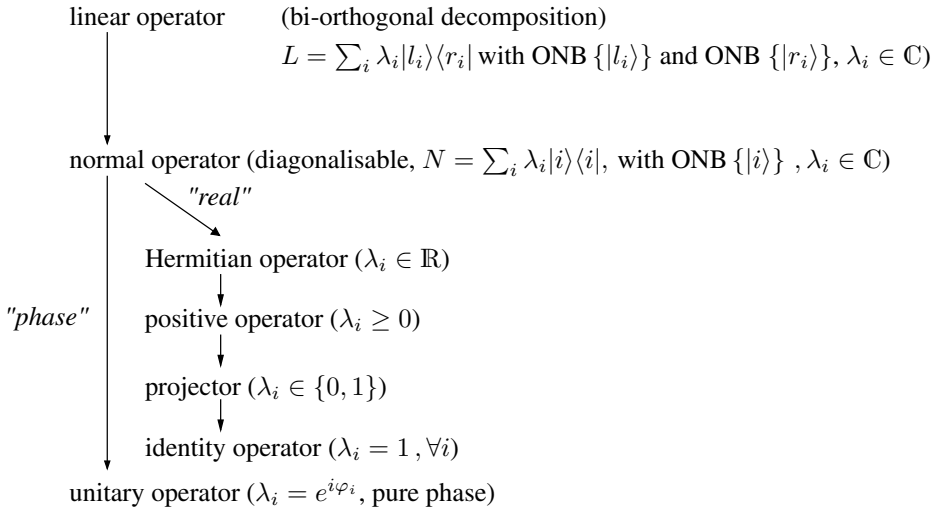

Figure 1.1: Hierarchy of operators. Characterisation of operators through their dyadic decomposition. $\rightarrow$ is in each case the direction of increasing specialisation. The eigenvalues are characterised in brackets ( ). The bi-orthogonal decomposition of a linear operator is derived in Sect. 13.3.3.

Functions of operators An *operator function* $f(N)$ is defined in terms of its expansion in a power series. For a normal operator N in the dyadic decomposition, it can be expressed in a simple way in terms of functions of the eigenvalues:

$$f(N) := \sum_i f(\lambda_i)|i\rangle\langle i| \quad \Rightarrow \quad f(N)|i\rangle = f(\lambda_i)|i\rangle \; . \tag{1.50}$$

$f(N)$ has the same eigenvectors $|i\rangle$ as N. We give an example which is formulated as a matrix representation with respect to the basis of the eigenvectors:

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = |0\rangle\langle 0| - |1\rangle\langle 1| \tag{1.51}$$

$$e^{\varphi\sigma_z} = e^{\varphi}|0\rangle\langle 0| + e^{-\varphi}|1\rangle\langle 1| = \begin{pmatrix} e^{\varphi} & 0 \\ 0 & e^{-\varphi} \end{pmatrix} \; . \tag{1.52}$$

1.1.4 Hermitian Operators

We follow the right-hand branch of the tree diagram in Fig. 1.1. A linear operator H on $\mathcal{H}_d$ is termed *Hermitian* or *self-adjoint* when it has the property $H^\dagger = H$. *Hermitian operators are special normal operators*. They play an important role in quantum mechanics, owing to their special properties: *Hermitian operators have a spectral decomposition with an ONB* $\{|i\rangle\}$

$$H = \sum_i r_i |i\rangle\langle i| \, , \quad r_i \in \mathbb{R} \tag{1.53}$$

and real eigenvalues r_i. In case of degeneracy, the eigenvectors can be chosen to be orthonormal, so that $\{|i\rangle\}$ forms an ONB. Eigenvectors belonging to different eigenvalues are orthogonal. This is often called the *spectral theorem*. Hermitian operators are also referred to as *observables*. The reason for this physical terminology will become clear later.

It follows immediately from Eq. (1.53) together with Eq. (1.35) that for an arbitrary vector $|\varphi\rangle$, the *expectation value* $\langle\varphi|H|\varphi\rangle$ is real. It is an important characteristic of Hermitian operators that the converse is also true: *the expectation value* $\langle\varphi|A|\varphi\rangle$ *is real for all vectors if and only if A is Hermitian.*

For a proof of the converse, we assume that for an operator A the mean value $\langle\chi|A|\chi\rangle$ is real for all vectors $|\chi\rangle$. For two arbitrary vectors $|\varphi\rangle$ and $|\psi\rangle$ from $\mathcal{H}$, the identity

$$\begin{aligned} 4\langle\varphi|A|\psi\rangle = \quad & \{(\langle\varphi| + \langle\psi|)A(|\varphi\rangle + |\psi\rangle) - (\langle\varphi| - \langle\psi|)A(|\varphi\rangle - |\psi\rangle)\} \\ + \quad & i[(\langle\varphi| + i\langle\psi|)A(|\varphi\rangle - i|\psi\rangle) - (\langle\varphi| - i\langle\psi|)A(|\varphi\rangle + i|\psi\rangle)] \end{aligned} \tag{1.54}$$

holds. If we exchange $|\varphi\rangle$ and $|\psi\rangle$ in this expression, then the part denoted by $\{\dots\}$ remains the same and the part $[\dots]$ changes its sign. Taking into account that all expectation values[2] are real, it then follows that $\langle\psi|A\varphi\rangle = \langle\varphi|A\psi\rangle^* = \langle A\psi|\varphi\rangle$. The operator A is thus Hermitian. It is notable that Eq. (1.54) contains on the right only expectation values, and on the left only a transition matrix element. *When all the expectation values of an Hermitian operator are known, then all the transition matrix elements are also known.*

The expectation value $\langle\varphi|A|\varphi\rangle$ is also called the *mean value*. Since their eigenvalues and mean values are real, Hermitian operators will play a special role in the theory of measurements (cf. Chap. 2).

Commuting Hermitian operators For these, the theorem on simultaneous diagonalisability holds (w/o.P.)[3]: *Two Hermitian operators (observables)* A *and* B *commute (*$[A, B]_- = 0$*) if and only if they have a common ONB* $\{|i\rangle\}$ *of eigenvectors.*

If the eigenvalue a of an observable A is degenerate, then the eigenvectors form a subspace which is at least two-dimensional. No associated eigenvector is therefore uniquely characterised by specifying a. If we consider only those eigenvectors of A within the subspace which are at the same time eigenvectors of an observable B which commutes with A and has eigenvalues b (intersecting sets), then a common eigenvector could be uniquely specified through this additional condition. We denote it by $|a, b\rangle$:

$$A|a, b\rangle = a|a, b\rangle, \quad B|a, b\rangle = b|a, b\rangle \; . \tag{1.55}$$

If only a subspace is determined in this way, then we continue and require that an eigenvector of A and B at the same time be an eigenvector of an observable C which commutes with A and B: $|a, b, c\rangle$. This procedure must be repeated until all degeneracies have been lifted. A set of observables which possesses exactly one common system of eigenvectors is called a *complete system of commuting observables*. Specifying the eigenvalues of all the operators determines a vector precisely. It is important that the procedure described in fact terminates. This is guaranteed by the following result (w/o.P.): *In every Hilbert space* $\mathcal{H}_\mathrm{d}$*, there exists a finite(!) complete set of operators which commute pairwise (functions of operators are not taken into consideration).* For the proof, we refer to the literature (cf. Sect. 1.4).

[2] According to Section 1.1.1, $\langle\varphi| + i\langle\psi|$ is the dual bra vector of $|\varphi\rangle - i|\psi\rangle$.

[3] w/o.P. means *without proof*

1.1.5 Unitary Operators

We first follow the left-hand branch of the operator-hierarchy tree in Fig. 1.1 and thereafter return to the right-hand branch. A linear operator U is called *unitary* when it has the property $U^\dagger = U^{-1}$. *Unitary operators are special normal operators. They have a spectral decomposition*

$$U = \sum_i e^{i\varphi_i} |i\rangle\langle i| \,, \quad \varphi_i \in \mathbb{R}, \tag{1.56}$$

with an ONB $\{|i\rangle\}$*, whereby, due to the defining equation, the eigenvalues are pure "phase factors". As with Hermitian operators, the eigenvectors span the entire space.* Eigenvectors with different eigenvalues are orthogonal. Eigenvectors with degenerate eigenvalues can be chosen to be orthogonal. As one can readily show, a linear operator is unitary precisely when each of its matrix representations is unitary. It follows immediately from the spectral decomposition that the operator function $U(t) = e^{iHt}$, $t \in \mathbb{R}$, is unitary if H is Hermitian. Furthermore, in this case:

$$U(t=0) = \mathbb{1} \tag{1.57}$$

$$U(t_2)U(t_1) = U(t_2 + t_1) \,. \tag{1.58}$$

Unitary equivalence and conservation of the norm Under combined unitary transformations of vectors and operators according to

$$|\varphi'\rangle = U|\varphi\rangle \quad A' = UAU^{-1} \,, \tag{1.59}$$

scalar products (in particular, the norm of a vector), eigenvalues and expectation values remain unchanged. *Conversely, a linear operator T, which conserves the norm on application to an arbitrary vector in $\mathcal{H}_d$,*

$$\|T\varphi\| = \|\varphi\| \,, \tag{1.60}$$

is a unitary operator: $T^\dagger = T^{-1}$. For the proof, we apply Eq. (1.7) and rewrite it using Eq. (1.60). For T, a unitarity relation holds:

$$\langle T\varphi|T\psi\rangle = \langle\varphi|\psi\rangle \,. \tag{1.61}$$

1.1.6 Positive Operators and Projection Operators

We wish to discuss some special cases of Hermitian operators (compare Fig. 1.1). A *positive operator* is defined by the fact that for an arbitrary vector $|\varphi\rangle$, the following inequality:

$$\langle\varphi|A|\varphi\rangle \geq 0 \quad \forall\, |\varphi\rangle \,, \tag{1.62}$$

holds, i.e. its expectation value is always real and non-negative. We can then write

$$A \geq 0 \,. \tag{1.63}$$

Furthermore, we define an inequality for operators:

$$A \geq B \Leftrightarrow (A - B) \geq 0 \,. \tag{1.64}$$

From the condition of positivity, it follows for the spectral decomposition that *every positive operator A is Hermitian, $A^\dagger = A$. It has the spectral decomposition*

$$A = \sum_i a_i |i\rangle\langle i|, \quad a_i \geq 0 \,. \tag{1.65}$$

with non-negative eigenvalues.

For an arbitrary linear operator A, $A^\dagger A$ is a positive operator. On the other hand, for each positive operator A there is a linear operator B such that A can be written in the form

$$A = B^\dagger B \,. \tag{1.66}$$

B is determined only up to unitary transformations ($B \to UB$). We can find B explicitly via the spectral decomposition of A (1.65) and an ONB $\{|\varphi_i\rangle\}$

$$B = \sum_i \sqrt{a_i}\, |\varphi_i\rangle\langle i| \,. \tag{1.67}$$

Substitution verifies (1.66).

A linear operator P is a *projection operator* (more precisely: an orthogonal projection operator) when it meets the following conditions:

$$P^2 = P \quad \textit{idempotent.} \tag{1.68}$$

$$P^\dagger = P \quad \text{Hermitian.} \tag{1.69}$$

It follows from these properties that

$$\langle v|P|v\rangle = \langle v|PP|v\rangle = \langle v|P^\dagger P|v\rangle = \|P|v\rangle\|^2 \geq 0 \,. \tag{1.70}$$

P is therefore a positive operator and fulfills

$$P = \sum_i p_i |i\rangle\langle i| \,; \quad p_i \geq 0 \tag{1.71}$$

with the ONB $\{|i\rangle\}$. Because it is idempotent, we furthermore have

$$P^2 = \sum_i p_i^2 |i\rangle\langle i| \,, \quad P = \sum_i p_i |i\rangle\langle i| \,, \tag{1.72}$$

and thus $p_i^2 = p_i$ or $p_i \in \{0, 1\}$. The projection operator P therefore assumes the form

$$P = \sum_{j \in I} |j\rangle\langle j|, \qquad I \leftrightarrow \text{subset of the ONB} \,. \tag{1.73}$$

P projects onto the subspace spanned by $\{|j\rangle\}$ with $j \in I$.

As a complement to Fig. 1.1, Fig. 1.2 shows retrospectively the "intersecting sets" of the different types of operators.

1.2 Liouville Operator Space

As we shall see in Chap. 2, in the special case of pure states, quantum-mechanical systems can be described by normalised vectors $|\psi\rangle$ in a Hilbert space $\mathcal{H}_{\rm d}$. In the general case of mixed quantum states, their description is accomplished using density operators (Chap. 4). All possible dynamical changes of state can be described in terms of linear transformations between density operators (Schrödinger representation). We will discuss this quite generally in Chap. 14. In preparation for this discussion, it is expedient to introduce the Liouville space $\mathbb{L}$ here. It is the space of the linear operators which act on the Hilbert space. We can restrict ourselves to a brief presentation, since the procedure is essentially a repetition of Sect. 1.1.

1.2.1 Scalar Product

The *Liouville space* $\mathbb{L}$ is a linear complex vector space whose elements $|A), |B), \ldots$ are the linear operators $A, B, \ldots$ which operate on a Hilbert space. One can readily verify that these linear operators in fact fulfill the axioms of a linear vector space. Later, we will leave off the brackets $|)$ in order to simplify our notation.

In this new notation, the dyadic decomposition (1.30) of an operator A in terms of the basis $\{|i\rangle\}$ on $\mathcal{H}_d$ has the form

$$|A) = \sum_{i,j=1}^{d} A_{ij} ||i\rangle\langle j|) \; . \tag{1.74}$$

The d^2 dyads $|i\rangle\langle j|$ in $\mathcal{H}_d$ make up the d^2 elements $||i\rangle\langle j|)$ of a basis of $\mathbb{L}$. For the dimensions of the spaces, we therefore have

$$\dim \mathbb{L} = (\dim \mathcal{H}_d)^2 \; . \tag{1.75}$$

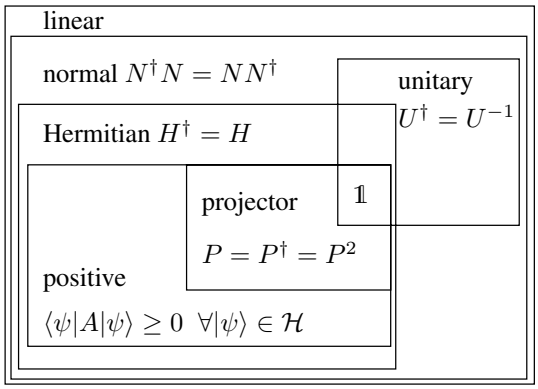

Figure 1.2: "Intersecting sets" of operator types. Note that for $\lambda_i \in \{1, -1\}$, special Hermitian operators can also be unitary and *vice versa*.

Of course, there are other basis sets besides the dyads in $\mathbb{L}$. The Liouville space $\mathbb{L}$ has a *scalar product* $(A|B)$. Formally, it has the same properties as the scalar product in the Hilbert space $\mathcal{H}_d$ (cf. Sect. 1.1.1). $(A|B)$ is a complex number and obeys the relations:

$$(A|B) = (B|A)^* \ , \ (A|c_1B_1 + c_2B_2) = c_1(A|B_1) + c_2(A|B_2) \ , \ (A|A) \geq 0 \ . \quad (1.76)$$

The operator basis Two operators A and B are called *orthogonal* if

$$(A|B) = 0 \quad (1.77)$$

holds, without either of the operators being the null operator. The triangle inequality (1.6) and an equation analogous to the parallelogram equation (1.8) also apply. Every operator $|A)$ can be decomposed in terms of an *orthonormal operator basis* $\{|Q_s), s = 1, \ldots, d^2\}$ of $\mathbb{L}$

$$(Q_s|Q_t) = \delta_{st}, \quad \sum_{s=1}^{d^2} |Q_s)(Q_s| = \mathbb{1} \ : \quad (1.78)$$

$$|A) = \sum_{s=1}^{d^2} |Q_s)(Q_s|A) \ . \quad (1.79)$$

The scalar product as trace Scalar products in $\mathbb{L}$ can be represented in various ways. *We will use the scalar product defined via the trace in $\mathcal{H}_d$*:

$$(A|B) := \mathrm{tr}[A^\dagger B] \ , \quad (1.80)$$

since in this case the Pauli spin operators, which are important for the simplest quantum systems, can be completed to form a basis (compare Sect. 3.1). The decomposition (1.79) –leaving off the vector brackets– takes on the form

$$A = \sum_{s=1}^{d^2} Q_s \, \mathrm{tr}[Q_s^\dagger A] \ . \quad (1.81)$$

The basis of the Liouville space generated from the dyads $|i\rangle\langle j|$ with $i, j = 1, \ldots, d$ is orthonormal with respect to the trace-scalar product (1.80)

$$\Big(|i\rangle\langle j| \Big| |i'\rangle\langle j'|\Big) = \delta_{ii'}\delta_{jj'} \ . \quad (1.82)$$

1.2.2 Superoperators

As might be presumed, we can define *linear* operators in the Liouville space itself, which map the elements of the space onto one another:

$$|A) \to \mathcal{S}|A) =: |\mathcal{S}A) =: \mathcal{S}(A) =: \mathcal{S}A \ . \quad (1.83)$$

These operators, which we write using italic symbols, are called *superoperators*. From the point of view of the Hilbert space $\mathcal{H}_d$, they map linear operators in a linear manner onto one another

$$A \to B = \mathcal{S}(A) \ . \quad (1.84)$$

Examples We give two examples of superoperators: For the superoperator $\mathcal{A}$

$$B \to \mathcal{A}(B) := ABA^{-1} \,, \tag{1.85}$$

linearity follows from the linearity of A. One can readily verify that

$$\mathcal{A}^{-1}(B) = A^{-1}BA \tag{1.86}$$

holds. An important superoperator for the description of the dynamic evolution of mixed states (compare Chap. 4) is the *Liouville operator* or *Liouvillian*, $\mathcal{L}$

$$A \to \mathcal{L}(A) := \frac{1}{\hbar}\,[H, A]_- \;. \tag{1.87}$$

In its application to physical problems, H in this expression is the Hamiltonian. The powers of $\mathcal{L}$ are written

$$\mathcal{L}^2(A) = \frac{1}{\hbar^2}\,\big[H, [H, A]_-\big]_- \;. \tag{1.88}$$

The concepts of adjoint, Hermitian, unitary and positive superoperators can be directly taken over from the corresponding definitions in Hilbert space.

1.3 The Elements of Probability Theory

As we have already emphasized, it is the central goal of quantum theory to make predictions concerning the probability of occurrence of measured values. To this end, it will be assumed that information about the state of the quantum object being measured is available. With this goal in mind, it is expedient to review briefly the basic concepts of probability theory here.

Predictions are conclusions drawn from the past and applied to the future. In classical physics, the reverse direction of conclusions plays a similarly important role. From the results of measurements, conclusions about the state of the object before the measurement are drawn. To what extent is this also possible for quantum systems? In the discussion of this question, Bayes' Theorem plays a very important role. We will sketch its proof after first presenting some preliminary considerations concerning conditional probabilities.

1.3.1 The Probability of Random Events

When a stochastic experiment is repeated, the result cannot be predicted. It is a *random event*. Such events could be for example the occurrence of an even or an odd number of dots when throwing dice, or the occurrence of a number greater than 2. Let $\{A_i;\, i = 1, \ldots, n\}$ be the number of such events. We introduce the following notation, in analogy to set theory:

$A_i \cap A_j \cap A_k$ is the event which consists of the simultaneous occurrence of the events A_i, A_j and A_k. In the case of dice, A_1 could be e.g. the event "even number of dots" and A_2 the event "number of dots > 4"; then $A_1 \cap A_2$ is the event "the six is thrown". $p(A_1 \cap A_2)$ is the probability that *both* A_1 *and also* A_2 occur (*joint probability*). We can also write

$$p(A_1, A_2) := p(A_1 \cap A_2) \,. \tag{1.89}$$

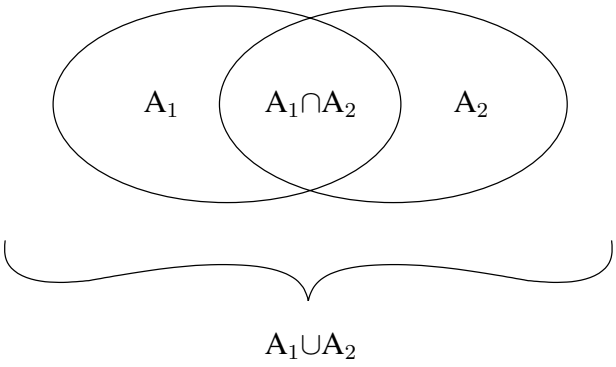

Figure 1.3: Set diagram for probabilities.

$A_i \cup A_j \cup A_k$ is the event consisting of the occurrence of *at least one* of the events A_i, A_j or A_k. For a number Z of dots, let $2 \leq Z \leq 4$ be the event A_1 and $3 \leq Z \leq 5$ be the event A_2. Then $A_1 \cup A_2$ is the event $2 \leq Z \leq 5$.

An impossible event is denoted by $\varnothing$ and a certain event by Ω. Two events A_i and A_j are called *exclusive events* when $A_i \cap A_j = \varnothing$. They cannot occur simultaneously.

Axioms With each random event A we associate a real number $p(A)$ with $0 \leq p(A) \leq 1$, which is called the *probability* of A, and which fulfills a series of axioms that we shall not list here. An example is given by Kolmogorov's axioms. We note only the *additivity axiom*: for pairwise exclusive random events $A_1, A_2, \ldots, A_n$ (i.e. $\mathrm{tr}(A_i \cap A_j) = 0$),

$$p(A_1 \cup A_2 \cup \ldots \cup A_n) = p(A_1) + p(A_2) + \ldots + p(A_n) \tag{1.90}$$

holds. When the events A_1 and A_2 are not exclusive, we find

$$p(A_1 \cup A_2) = p(A_1) + p(A_2) - p(A_1 \cap A_2)\ . \tag{1.91}$$

The set diagram in Fig. 1.3 gives an intuitive picture of this relation. For thrown dice, let $Z \leq 2$ be event A_1 and $Z \geq 4$ be event A_2; then the probability that either A_1 or A_2 occurs is $p(A_1 \cup A_2) = \frac{2}{6} + \frac{3}{6} = \frac{5}{6}$.

Frequency interpretation In order to make the axiom clear, we used the example of throwing dice. In fact, this axiom, like all mathematical axioms, requires no physical interpretation. $p(A)$ is defined by the axioms themselves. When applied to physical events, probability is usually interpreted as the *relative frequency*:

$$p(A) := \lim_{N \to \infty} \frac{N(A)}{N} \tag{1.92}$$

Here, $N(A)$ is the absolute frequency of occurrence of the event A in a total number N of attempts. This physical interpretation is not without problems. For a finite number N, it can be taken as an estimate of $p(A)$.

1.3.2 Conditional Probability and Bayes' Theorem

We extend the concept of probability. The *conditional probability* $p(A|B)$ of an event A is the probability of occurrence of A *under the condition that* another event B, which itself has the probability $p(B)$, has already occurred. We define:

$$p(A|B) := \frac{p(A \cap B)}{p(B)} \ . \tag{1.93}$$

Resolution of this expression leads to the plausible equation for the probability $p(A \cap B)$ for the occurrence of both A and B:

$$p(A \cap B) = p(A|B) \cdot p(B) \ . \tag{1.94}$$

As an example, we consider two urns. The urn U_1 contains 3 white and 3 black balls; urn U_2 contains 2 white and 4 black balls. From each of the urns, balls are picked with the same probability $p(U_1) = p(U_2) = \frac{1}{2}$. The probability of being picked is the same for every ball, i.e. $\frac{1}{12}$. The probability of picking a ball both from U_1 and also that the ball picked be white is given by $p(w \cap U_1) = \frac{3}{12} = \frac{1}{4}$. The conditional probability $p(w|U_1)$ of getting a white ball when picking from urn U_1 is given according to Eq. (1.93) by

$$p(w|U_1) = \frac{p(w \cap U_1)}{p(U_1)} = \frac{2}{4} = \frac{1}{2} \ . \tag{1.95}$$

This follows intuitively directly from the description of the randomness of the situation. Analogously, one finds $p(w|U_2) = \frac{1}{3}$.

Independence Two random events A and B are called *independent* events when the occurrence of the one event has no influence on the probability of occurrence of the other,

$$p(A|B) = p(A) \ . \tag{1.96}$$

In this case, it follows with (1.93) that

$$p(A \cap B) = p(A)p(B) \ . \tag{1.97}$$

From this, it must be distinguished whether the events A and B are exclusive (mutually contradictory), $A \cap B = \varnothing$. In that case, we have $p(A|B) = 0$.

Total probability The certain event Ω can be represented as the sum of n pairwise exclusive random events A_i, $(A_i \cap A_j = \varnothing, \quad \forall i \neq j)$:

$$\Omega = A_1 \cup A_2 \cup \ldots \cup A_n; \ A_i \cap A_j = \varnothing, \quad \forall \, i \neq j \ . \tag{1.98}$$

For an arbitrary random event B, we then find $B = (A_1 \cap B) \cup (A_2 \cap B) \cup \ldots \cup (A_n \cap B)$. From the additivity axiom (1.90) it follows that

$$p(B) = \sum_{i=1}^{n} p(B \cap A_i) \ , \tag{1.99}$$

and with Eq. (1.94) we obtain the *addition theorem for probabilities* or the total probability

$$p(B) = \sum_{i=1}^{n} p(B|A_i)p(A_i) \ . \tag{1.100}$$

We will give an example in the next section.

Bayes' theorem With $p(A \cap B) = p(B \cap A)$, Eq. (1.94) for random events A_i leads to

$$p(A|B)p(B) = p(B|A)p(A) \ . \tag{1.101}$$

Under the assumption of pairwise exclusivity and completeness (1.98), we obtain with Eq. (1.100) the fundamental Bayes' Theorem

$$p(A_i|B) = \frac{p(B|A_i)p(A_i)}{\sum_{j=1}^{n} p(B|A_j)p(A_j)} \ . \tag{1.102}$$

The denominator guarantees the normalisation $\sum_i p(A_i|B) = 1$, which ensures that one of the events A_i must occur.

Bayes' theorem can be interpreted as follows: let the probabilities $p(A_i)$ and the conditional probabilities $p(B|A_i)$ for B given A_i be known for a certain situation. Then equation (1.102) permits the computation of the conditional probability $p(A_i|B)$ for A_i given B. If the event B occurs *after* event A_i, then $p(A_i|B)$ answers the following question: if B occurs, what was the probability that A_i had already occurred? This conclusion for A_i given B applies in the reverse direction to that of $p(B|A_i)$. This demonstrates the significance of the theorem.

We give an example, again based on picking balls from urns. Let us presume the existence of three urns of type I with 2 white and 6 black balls in each one, and of one urn of type II with 1 white and 7 black balls. The probability of choosing any one of the urns from which to pick a ball is the same. Result B means that a white ball is picked. The event A_1 (or A_2) means that a ball is picked from an urn of type I (or of type II). Then we have the following probabilities: $p(A_1) = \frac{3}{4}, p(A_2) = \frac{1}{4}, p(B|A_1) = \frac{1}{4}, p(B|A_2) = \frac{1}{8}$. The probability that a white ball picked comes from an urn of type I is, according to Bayes' theorem, given by $p(A_1|B) = \frac{6}{7} = 0.86$ and is thus greater than $p(A_1)$. A white ball comes from the urn of type II with a probability $p(A_2|B) = \frac{1}{7} = 0.14$, which is less than $p(A_2)$. The choice of the type of urn is made with the *a priori* probabilities $p(A_i)$. If a white ball was picked, one can make an inference about which urn it originated from. For this *inference* there is in general only a probability statement which is given by $p(A_i|B)$. If the urn of type II contained no white balls, the inference could be made with certainty $(p(A_1|B) = 1)$ that a white ball was picked from an urn of type I.

The following explanation of Bayes' theorem can also be helpful: we consider the special case that all the $p(A_i)$ are equal. The event B is to be predicted. The event A_k for which the subsequent occurrence of B is most probable (i.e. $p(B|A_k) = \max$) has also previously occurred with maximum probability, $p(A_k|B) = \max$.

Bayes' assumption This is not to be confused with Bayes' theorem. If there is no reason to assume that an event A_i is particularly preferred by the situation, it can be reasonable to make *Bayes' assumption* that all the *a priori* probabilities are equal,

$$p(A_1) = p(A_2) = \ldots = p(A_n) \; . \tag{1.103}$$

This assumption is not the same as Bayes' theorem. After the occurrence of B, this assumption is replaced by the probabilities $p(A_i|B)$ from Eq. (1.102). The probabilities can be estimated in this way.

1.3.3 Random Quantities

A *random variable* X is given by association of numbers x to the corresponding random events. Throws of dice can be used as an illustrative example. A discrete random quantity X is determined by the values $x_1, x_2, \ldots, x_n$ and the probabilities $p(x_1), p(x_2), \ldots, p(x_n)$ with which the values occur ($\sum_{i=1}^{n} p_i = 1$). The generalisation to an enumerable infinity of values x_i and to a continuous variable x is usually not problematic.

Important values for the characterisation of a random variable X are the *expectation value* or the *mean value*

$$\langle X \rangle := \sum_i p_i x_i \tag{1.104}$$

and the *dispersion* or the *mean square deviation*

$$\mathrm{var}(X) = (\Delta X)^2 := \langle X^2 \rangle - \langle X \rangle^2 = \langle (X - \langle X \rangle)^2 \rangle \; , \tag{1.105}$$

which is also called the *variance* var(X). The *standard deviation* $\Delta X = \sqrt{\mathrm{var}(X)}$ indicates how widely the random variable is distributed around its mean value. In quantum mechanics, ΔX is also referred to as the *uncertainty* of X.

1.4 Complementary Topics and Further Reading

- Most textbooks of quantum mechanics contain a summary of the mathematical fundamentals. We mention in particular the following books: [Sak 85], [Ish 95], [Bal 98], [Gri 02], [CDL 05].

- A detailed treatment of Hilbert space with reference to quantum mechanics can be found in [Jor 69].

- The bra space as the vector space of all linear continuous functionals in a vector space V (also called the dual space V^*): [FK 98, Chaps. 2.8 and 4.2].

- A collection of references for Sect. 1.3: [Ish 95], [NC 00].

1.5 Problems for Chapter 1

Prob. 1.1 [for Sect. 1.1]: Prove the relations (1.5), (1.6), (1.7), (1.8), (1.24), (1.25), (1.27), (1.28), (1.30), (1.35), (1.50), (1.61).

Prob. 1.2 [for Sect. 1.1]: Prove that a linear operator which acts on a finite-dimensional complex vector space has at least one eigenvector and one eigenvalue.

Prob. 1.3 [for Sect. 1.1]: Give several examples of a basis set of $\mathcal{H}_3$.

Prob. 1.4 [for Sect. 1.1]: Let $\{|i\rangle, i = 1, \ldots, d\}$ be an ONB. Prove that Parseval's identity

$$\|\varphi\|^2 = \sum_{i=1}^{n} |\langle\varphi|i\rangle|^2 \tag{1.106}$$

holds for all vectors $|\varphi\rangle \in \mathcal{H}_2$.

Prob. 1.5 [for Sect. 1.1]: Show that the matrix which corresponds to the operator product AB is equal to the product of the matrices of A and B.

Prob. 1.6 [for Sect. 1.1]: Show that every linear operator C can be written in the form

$$C = R + iI \tag{1.107}$$

with Hermitian operators R and I. Consider the analogy: linear operator $\leftrightarrow$ complex number; Hermitian operator $\leftrightarrow$ real number.

Prob. 1.7 [for Sect. 1.1]: Show that the determinant of a unitary matrix is ± 1.

Prob. 1.8 [for Sect. 1.1]: Show that for two unitary $n \times n$ matrices U_1 and U_2, the matrix $\begin{pmatrix} U_1 & 0 \\ 0 & U_2 \end{pmatrix}$ is also unitary.

Prob. 1.9 [for Sect. 1.1]: Does the projection operator $P = |u\rangle\langle u|$ have an inverse?

Prob. 1.10 [for Sect. 1.1]:

a) The operator A is known to be diagonalisable. How can its spectral representation be found?

b) Are the Pauli operators $\sigma_x = |0\rangle\langle 1| + |1\rangle\langle 0|, \sigma_y = -i|0\rangle\langle 1| + i|1\rangle\langle 0|$, and $\sigma_z = |0\rangle\langle 0| - |1\rangle\langle 1|$ diagonalisable? Find their spectral representation.

Prob. 1.11 [for Sect. 1.1]: Give an example of a normal operator which is neither Hermitian nor unitary.

Prob. 1.12 [for Sect. 1.2]: Confirm that the relation

$$\mathrm{tr}[C(\mathcal{A}B)] = \mathrm{tr}[(\mathcal{A}^{-1}C)B] \tag{1.108}$$

holds for the superoperator A defined in Eq. (1.85).

Prob. 1.13 [for Sect. 1.2]: Let H be a Hermitian operator which obeys the eigenvalue equation

$$H|e_i\rangle = E_i|e_i\rangle \ . \tag{1.109}$$

Find the eigenvectors and the eigenvalues of the Liouville operator $\mathcal{L}$ from Eq. (1.87).

Prob. 1.14 [for Sect. 1.2]: Show that the Liouville operator from Eq. (1.87) has the matrix representation

$$\mathcal{L}_{ij,i'j'} = \frac{1}{\hbar}(H_{ii'}\delta_{j'j} - \delta_{ii'}H_{j'j}) \ . \tag{1.110}$$

Prob. 1.15 [for Sect. 1.2]: Prove the following relation by referring to the definition of the Liouville operator $\mathcal{L}$:

$$e^{c\mathcal{L}}A = e^{\frac{c}{\hbar}H}Ae^{-\frac{c}{\hbar}H} \ . \tag{1.111}$$

Prob. 1.16 [for Sect. 1.2]: Describe some situations which can be used to give an intuitive understanding of conditional probability, of the total-probability theorem, or of Bayes' theorem.